

Random walks which prefer unvisited edges. Exploring high girth even degree expanders in linear time

Petra Berenbrink* Colin Cooper† Tom Friedetzky‡

March 2, 2013

Abstract

In this paper, we consider a modified random walk which uses unvisited edges whenever possible, and makes a simple random walk otherwise. We call such a walk an *edge-process* (or *E-process*). We assume there is a rule \mathcal{A} , which tells the walk which unvisited edge to use whenever there are several unvisited edges. In the simplest case, \mathcal{A} is a uniform random choice over unvisited edges incident with the current walk position. However we do not exclude arbitrary choices of rule \mathcal{A} . For example, the rule could be determined on-line by an adversary, or could vary from vertex to vertex.

For the class of connected, even degree graphs G of constant maximum degree, we characterize the vertex cover time of the E -process in terms of the edge expansion rate of G , as measured by eigenvalue gap $1 - \lambda_{\max}$ of the transition matrix of a simple random walk on G .

A vertex v is ℓ -good, if any even degree subgraph containing all edges incident with v contains at least ℓ vertices. A graph G is ℓ -good, if every vertex has the ℓ -good property.

In particular, for even degree expander graphs, of bounded maximum degree, we have the following result. Let G be an n vertex ℓ -good expander graph. Any E -process on G has cover time

$$C_G(E\text{-process}) = O\left(n + \frac{n \log n}{\ell}\right).$$

This result is independent of the rule \mathcal{A} used to select the order of the unvisited edges, which can be chosen on-line by an adversary.

With high probability random r -regular graphs, $r \geq 4$ even, are expanders for which $\ell = \Omega(\log n)$. Thus, for almost all such graphs, the vertex cover time of the E -process is $\Theta(n)$. This improves the vertex cover time of such graphs by a factor of $\log n$, compared to the $\Omega(n \log n)$ cover time of *any* weighted random walk.

*School of Computing Science, Simon Fraser University, Burnaby, Canada

†Department of Informatics, King's College, London, U.K.

‡School of Engineering and Computing Sciences, Durham University, Durham, U.K.

1 Introduction

In a simple random walk on a graph, at each step a particle moves from its current vertex position to a neighbouring vertex chosen uniformly at random. Formally, a *simple* random walk $\mathcal{W}_v = (\mathcal{W}_v(t), t = 0, 1, \dots)$ is defined as follows: $\mathcal{W}_v(0) = v$ and given $x = \mathcal{W}_v(t)$, $y = \mathcal{W}_v(t + 1)$ is a randomly chosen neighbour of x .

In this paper, we consider a modified walk which uses unvisited edges whenever possible, and makes a simple random walk otherwise. We call such a walk an *edge-process* (or *E-process*). At each step the edge-process makes a transition to a neighbour of the currently occupied vertex as follows:

If there are *unvisited edges* incident with the current vertex pick one and make a transition along this edge.

If there are no unvisited edges incident with the current vertex, move to a random neighbour using a simple random walk.

If we wish, can we assume there is a rule \mathcal{A} , which tells the walk which unvisited edge to use whenever there is a choice. In the simplest case, this is a uniform random choice over unvisited edges incident with the current walk position. However we do not exclude arbitrary choices of rule \mathcal{A} . For example, the rule could be deterministic or decided on-line by an adversary, or could vary from vertex to vertex.

The *E-process* seems particularly adapted to searching in a physical environment, where edges can easily be marked as visited. Imagine walking in a labyrinth, and marking the entries and exits of the edges taken with a piece of chalk. Whenever all exits are marked, walk randomly.

For any process which explores a graph G by walking from vertex to vertex, the *vertex cover time*, C_G , is defined as follows. For $v \in V$, let C_v be the expected time taken for a walk W on G starting at v , to visit every vertex of G . The vertex cover time is defined as $C_G = \max_{v \in V} C_v$. It was shown by Feige [9], that for any connected n -vertex graph G , the cover time of a simple random walk satisfies $C_G \geq (1 - o(1))n \log n$. In fact, any weighted reversible random walk has a lower bound on the cover time of $C_G = \Omega(n \log n)$. Thus no reversible random walk can have an $o(n \log n)$ cover time. A proof of the $\Omega(n \log n)$ lower bound on the cover time of weighted random walks, due to T. Radzik [15], is given in Section 3.1.

One random process similar to the *E-process*, is the *Random Walk with Choice*, ($\text{RWC}(d)$), of Avin and Krishnamachari [2]. The process $\text{RWC}(d)$ selects d neighbours uniformly at random at each step, and moves to the least visited vertex among them. The paper [2] makes an experimental study of the process $\text{RWC}(d)$ on geometric random graphs, and the toroidal grid, and finds reductions in cover time, and improved concentration of experimental results. Recently a special case of the *E-process* has been studied by Orenshtein and Shinkar [14] in the context of edge cover times. In [14], the next unvisited edge is chosen u.a.r. For a further discussion on edge cover time see below.

In the context of deterministic walks, the *E-process* has similarities with the rotor-router, or Propp machine model; see [7] for an introduction to this topic. The analysis of both processes depends on the underlying Eulerian properties of the graph. In the case of the rotor-router process, the graph is turned into an Eulerian digraph by replacing each edge with a pair of oppositely directed edges. The vertex cover time of the rotor-router model is $O(mD)$, where m is the number of edges of G , and D is the diameter, see [17].

The class of graphs we consider are connected, even degree graphs G of constant maximum degree $\Delta(G)$. We define a local expansion property of vertices. We say a vertex v is ℓ -good, if any even

degree subgraph containing all edges incident with v contains at least ℓ vertices. A graph G is ℓ -good, if every vertex has the ℓ -good property. We characterize the cover time of the E -process in terms of the edge expansion rate of G , as measured by eigenvalue gap $1 - \lambda_{\max}$ of the transition matrix of a simple random walk on G . A general statement of our result is the following theorem.

Theorem 1 *Let G be a connected n vertex even degree graph, with finite maximum degree, and the additional property that G is ℓ -good. Then, **any** E -process on G has cover time*

$$C_G(E\text{-process}) = O\left(n + \frac{n \log n}{\ell(1 - \lambda_{\max})}\right).$$

We briefly list a series of remarks and corollaries which arise from Theorem 1

- i) The upper bound on the cover time given in Theorem 1 is independent of the rule \mathcal{A} used to select unvisited edges, even if this choice is decided on-line by an adversary.
- ii) For expander graphs, which have positive constant eigenvalue gap, Theorem 1 becomes

$$C_G(E\text{-process}) = O\left(n + \frac{n \log n}{\ell}\right). \quad (1)$$

In particular, for ℓ -good even degree expanders where $\ell = \Omega(\log n)$, the E -process covers the graph in $\Theta(n)$ steps. As any walk-based process must take n steps to visit every vertex, the order of our result is best possible.

- iii) Examples of ℓ -good graphs where $\ell = \Omega(\log n)$ include random r -regular graphs, for which we have the following corollary.

Corollary 2 *Let $r \geq 4$ even. Let \mathcal{G}_r denote the class of random r -regular graphs. Let G be sampled uniformly at random from \mathcal{G}_r , then with high probability $C_G(E\text{-process}) = O(n)$.*

See Section 3 for the proof of this. Other examples of ℓ -good graphs are random graphs with fixed degree sequence \mathbf{d} , and all vertices of even degree at least 4, (ii) algebraically constructed even degree expanders of logarithmic girth, see [12].

- iv) The lower bound on the cover time of G by *any* weighted reversible random walk is $\Omega(n \log n)$. (See Section 3.1 for a proof of this result). For expanders, the comparable cover time is given by (1). Up to $\ell = \log n$, this gives a speed up of $\Omega(\log n / \ell)$ compared to *any* random walk.
- v) In Section 3.3 we give some experimental results on the performance of the E -process. Simulations suggest that for even degree random regular graphs, the cover time of the E -process is bounded (asymptotically) by the number of edges m in the graph (see Figure 1).

Could we expect an $O(n)$ cover time for the E -process on odd degree expanders? Experimentally, we find that this is not the case (see Figure 1).

- vi) A practical consequence of Theorem 1, is that, in order to build ‘easy to search’ networks, we should ensure all vertices have even degree and few short cycles. Examples of such constructions, based on even degree random r -regular graphs, are the SWAN P2P network of [4] based on switches, and the flip based P2P network of [13]. Properties of these networks such as connectivity, diameter and mixing-rate were studied in (e.g.) [5], [6], [8].

We also make some observations on edge cover time of the E -process (see i,ii below), and on the relationship between the E -process and Propp machines (items iii-v).

- i) In general upper bounds on the edge cover time of the E -process depend on the number of short cycles. The girth g of a graph G is the minimum length cycle in G . It can be shown that the E -process will cover all edges of a connected even degree graph in $O(|E| + n \log n / (1 - \lambda_2)^2 g)$. This bound can be improved if the number of short cycles can be upper bounded. As an example, for even degree random regular graphs, the (**whp**) upper bound on the edge cover time is $O(n\omega)$, where $\omega \rightarrow \infty$ arbitrarily slowly.
- ii) The result of [14] gives a bound for edge cover time of r -regular graphs of $O(|E| + n \log n / (1 - \lambda_{\max}))$. This is at best $O(n \log n)$ for sparse graphs, but is tight for expanders provided the number of edges $|E| = \Omega(n \log n)$. This result differs qualitatively from Theorem 1 which treats vertex cover time of constant degree expanders ($|E| = cn$, c constant).
- iii) Suppose it is the case that the edges of a graph G can be distinguished as unvisited in each direction by the E -process; i.e. a first visit (x, y) and a first visit (y, x) are regarded as distinct. This converts G into an Eulerian digraph, so that the even degree restriction is no longer necessary, and Theorem 1 now holds for all connected graphs of bounded degree.
- iv) Suppose the edges of the graph can be marked as unvisited in each direction. Then the ordering of the (directed) unvisited edges at each vertex made by the rule \mathcal{A} is a rotor order for a rotor-router (Propp machine). The E -process acts as a hybrid of a Propp machine and a random walk, the algorithm being: *Use the rotor once at each vertex and then walk randomly. Any rotor order will do.* The power of the adversary is to set the rotor order.
- v) In some rotor-router models an adversary can force a cover time of $\Omega(m \log m)$ on connected m edge graphs (see [3] for details). This phenomena partially arises because the adversary can make the walk retrace visited edges, even when unvisited edges are present at a vertex. In the E -process the adversary is less strong, and only has power to select the next unvisited edge used by the process. All transitions over visited edges are chosen randomly. Thus when $m = \Theta(n)$, and $\ell = \Omega(\log n)$, the E -process has cover time $\Theta(n)$, as compared to $\Theta(n \log n)$ cover time in the aforementioned adversarial rotor-router model.

1.1 Random walk properties

Let $G = (V, E)$ denote a connected graph, $|V| = n$, $|E| = m$, and let $d(v)$ be the degree of a vertex v . A *simple random walk* \mathcal{W}_u , $u \in V$, on graph G is a Markov chain modeled by a particle moving from vertex to vertex according to the following rule. The probability of transition from vertex v to vertex w is equal to $1/d(v)$, if w is a neighbour of v , and 0 otherwise. The walk \mathcal{W}_u starts from vertex u at $t = 0$. Denote by $\mathcal{W}(t)$ the vertex reached at step t ; $\mathcal{W}(0) = u$.

Let P be the transition matrix of a simple random walk on a graph G . Thus $P_{i,j} = 1/d(i)$ if and only if there is an edge between i and j in G . Let $P_u^{(t)}(v) = \mathbf{Pr}(\mathcal{W}_u(t) = v)$ be the t -step transition probability. We assume the random walk \mathcal{W}_u on G is ergodic with stationary distribution π , where $\pi_v = d(v)/(2m)$. If this is not the case, e.g. G is bipartite, then the walk can be made ergodic, by making it *lazy*. A random walk is *lazy*, if it moves from v to one of its neighbours w with probability $1/(2d(v))$, and stays where it is (at vertex v) with probability $1/2$.

Let $1, \lambda_2, \dots, \lambda_n$, be the eigenvalues of P , and let $\lambda_{\max} = \min(|\lambda_2|, |\lambda_n|)$. We henceforth assume that $\lambda_2 = \lambda_{\max}$ which can be achieved by making the chain lazy. This has no significant effect on our analysis.

The convergence to stationarity of a simple random walk is bounded by

$$|P_u^{(t)}(x) - \pi_x| \leq (\pi_x/\pi_u)^{1/2} \lambda_{\max}^t. \quad (2)$$

Visits to a Single Vertex For a random walk starting from vertex u , let H_v be the number of steps taken to reach vertex v , and let $\mathbf{E}_u(H_v)$ be the expected value of H_v ; the expected hitting time of v starting from u . If the distribution of the random walk at some step is $\rho = (\rho(u), u \in V)$, we can similarly define the hitting time from starting distribution ρ as $\mathbf{E}_\rho(H_v) = \sum_{u \in V} \rho(u) \mathbf{E}_u(H_v)$.

For a random walk starting at a vertex chosen from the stationary distribution π , let $\mathbf{E}_\pi(H_v)$ denote the expected hitting time of vertex v from stationarity. The quantity $\mathbf{E}_\pi(H_v)$ can be expressed in the following way, (see e.g. [1], Chapter 2)

$$\mathbf{E}_\pi(H_v) = Z_{vv}/\pi_v, \quad (3)$$

where

$$Z_{vv} = \sum_{t=0}^{\infty} (P_v^{(t)}(v) - \pi_v). \quad (4)$$

Using (2), we can bound the value of $\mathbf{E}_\pi(H_v)$ as follows.

Lemma 3

$$\mathbf{E}_\pi(H_v) \leq \frac{1}{(1 - \lambda_{\max})\pi_v}. \quad (5)$$

Proof Using (2) with $x = u = v$, then

$$|P_v^t(v) - \pi_v| \leq (\lambda_{\max})^t,$$

and

$$Z_{vv} = \sum_{t \geq 0} (P_v^t(v) - \pi_v) \leq \sum_{t \geq 0} (\lambda_{\max})^t = \frac{1}{1 - \lambda_{\max}}.$$

□

Let T_G be the mixing time of a graph G , such that, for $t \geq T_G$,

$$\max_{u, x \in V} |P_u^{(t)}(x) - \pi_x| = O\left(\frac{1}{n^3}\right). \quad (6)$$

Let $\mathcal{A}_t(v) = \mathcal{A}_{t,u}(v)$ denote the event that \mathcal{W}_u does not visit vertex v in steps $0, \dots, t$. Lemma 4 gives a bound for $\Pr(\mathcal{A}_t(v))$ in terms of $\mathbf{E}_\pi(H_v)$ and the mixing time T .

Lemma 4 *Let T_G be the mixing time of a random walk \mathcal{W}_u on G satisfying (6). Then*

$$\Pr(\mathcal{A}_t(v)) \leq e^{-\lfloor t/(T_G + 3\mathbf{E}_\pi(H_v)) \rfloor}.$$

Proof Let $\rho = (\rho_w)$ be the distribution of \mathcal{W}_u on G after $T = T_G$ steps, where $\rho_w = P_u^{(T)}(w)$. Let $\mathbf{E}_\rho(H_v)$ be the expected time to hit v starting from ρ . As T satisfies (6), and $\pi_x = \Omega(1/n^2)$ for any connected graph, then $\rho_w = (1 + o(1))\pi_w$. It follows that

$$\mathbf{E}_\rho(H_v) = (1 + o(1))\mathbf{E}_\pi(H_v). \quad (7)$$

Let $H_v(\rho)$ be the time to hit v starting from ρ , then

$$\Pr[H_v(\rho) \geq 3\mathbf{E}_\pi(H_v)] \leq \frac{1}{e}.$$

Let $\tau = T + 3\mathbf{E}_\pi(H_v)$. By considering the process \mathcal{W}_u at $\mathcal{W}(0) = u, \mathcal{W}(\tau), \mathcal{W}(2\tau), \dots, \mathcal{W}(\lfloor t/\tau \rfloor \tau)$ we obtain

$$\Pr(\mathcal{A}_t(v)) \leq e^{-\lfloor t/\tau \rfloor}.$$

□

Visits to Vertex Sets We can extend the results presented above to any nonempty subset S of vertices in the following way. From G we obtain a (multi)-graph $\Gamma = \Gamma_S$ by contracting S to a single vertex γ . Note that we retain multiple edges and loops in Γ_S , so that $d(S) = d(\gamma)$, and $|E(\Gamma)| = |E(G)| = m$. Let $\hat{\pi}$ be the stationary distribution of a random walk on Γ . If $v \notin S$ then $\hat{\pi}_v = \pi_v$, and $\hat{\pi}_\gamma = \pi_S \equiv \sum_{x \in S} \pi_x$.

For $u \notin S$ let \mathcal{W}_u be a walk starting from u in G , and let $\widehat{\mathcal{W}}_u$ be the equivalent walk starting in Γ . Provided \mathcal{W}_u does not visit S in t steps, (the event $\mathcal{A}_t(S, G)$), then $\widehat{\mathcal{W}}_u$ does not visit γ (the event $\mathcal{A}_t(\gamma, \Gamma)$), and the walks have the same transition probabilities. Thus,

$$\Pr(\mathcal{A}_t(S, G)) = \Pr(\mathcal{A}_t(\gamma, \Gamma)),$$

and

$$\mathbf{E}_\pi(H_S) = \mathbf{E}_{\hat{\pi}}(H_\gamma). \quad (8)$$

It is a known result that contracting vertex sets increases the eigenvalue gap. (For a proof see e.g. [1] Chapter 3, Corollary 27.) Thus

$$1 - \lambda_{\max}(G) \leq 1 - \lambda_{\max}(\Gamma).$$

In our proofs, we will always choose a mixing time T in (6) satisfying both $T \geq T_G$, and $T \geq T_\Gamma$. It follows that, using this mixing time T , the results of Lemma 3, and Lemma 4 apply equally to Γ , and to G . Thus e.g.

Corollary 5 *Let $G = (V, E)$, let $|E| = m$. Let $S \subseteq V$, and let $d(S)$ be the degree of S . Then $\mathbf{E}_\pi H_S$, the expected hitting time of S from stationarity satisfies*

$$\mathbf{E}_\pi H_S \leq \frac{2m}{d(S)(1 - \lambda_{\max}(G))}.$$

2 Proof of main result

2.1 Properties of the edge-process

It is helpful to think of the progress of the E -process as a re-colouring of the edges of the graph G . We consider unvisited edges as coloured blue, and explored edges as coloured red. Let $X(t)$ be the position at step t of a particle moving according to an E -process.

Initially, the particle is at $X(0) = u$, the start vertex, and all edges of the graph G are coloured blue (unvisited). Given $X(t) = v$, $X(t+1)$ is chosen as follows. If all edges incident with v are red (previously visited) the walk chooses $X(t+1)$ u.a.r. from $N(v)$. If however, there are any blue (unvisited) edges incident with v , then we pick a blue edge (v, w) according to the rule \mathcal{A} . The walk then moves to $X(t+1) = w$, and re-colours the edge (v, w) red (visited). We assume that the edge (v, w) is re-coloured red at the start of step $t+1$, the instant at which the walk arrives at w . Thus we regard the transition (v, w) as being along a blue edge.

At each t the next transition is either along a blue or a red edge. We speak of the sequence of these edge transitions as the blue (sub)-walk and the red (sub)-walk. The walk thus defines red and blue phases which are maximal sequences of edge transitions when the walk is the given colour. For any vertex v , and step t , the blue (resp. red) degree of v is the number of blue (resp. red) edges incident with v at the start of step t .

Observation 6 *Assume all vertices of G are of even degree. Then a blue phase of the E -process which starts at a vertex v (at some step t), must end at v (at some step $t + \tau$).*

Proof This follows from a simple parity argument. The first blue phase starts at $t = 0$, at the start vertex u . At $t = 0$ every vertex has even blue degree. Suppose that at step t we have $X(t) = w$, where $w \neq u$. Inductively every vertex, apart from the start vertex u and the current position w have even blue degree, whereas the blue degree of u and w is odd, and hence greater than zero. The particle can thus exit w along a blue edge. When the particle leaves $w = X(t)$ making the transition $(X(t), X(t+1))$, then the blue degree of $w = X(t)$ becomes even. If $X(t+1) = u$, then the degree of u is even and the particle has returned to the start. If $X(t+1) \neq u$, then the blue degree of $X(t+1)$ and u is odd.

If the particle returns to u at step t , and the blue degree of u is zero, then the blue phase at u is completed at (the start of) step t . The particle now leaves u along a red edge $(u, v) = (X(t), X(t+1))$, and this is the beginning of a red phase. Inductively, the blue degree of v is even when the particle arrives at v . If v has blue edges incident with it, then a blue phase begins. Otherwise the red phase continues. \square

Note that it is possible for all edges incident with a vertex v to be coloured red by transitions made during the blue sub-walk, and that v has not been visited by a red walk.

Let $G[S]$ denote the subgraph of G induced by the set of vertices $S \subseteq V$. The following summarizes the consequences of Observation 6.

Observation 7 *Assume vertex v is unvisited at step t , and that the E -process is in a red phase.*

1. *All edges incident with v are blue at step t .*
2. *The blue degree of all vertices at step t is even.*

3. Let S_v^* be the maximal blue (unvisited), edge induced subgraph obtained by fanning out in a breadth first manner from v using only blue edges. Let U^* be the vertex set of S_v^* . Then

- (a) The degree of v in S_v^* is $d(v)$, the degree of vertex v in G . All vertices of S_v^* have positive even degree.
- (b) All edges between S_v^* and $G \setminus U^*$ are red.
- (c) $G[U^*]$ may induce red edges, but these are not part of S_v^* .

In the simplest case S_v^* consists of $d(v)/2$ blue cycles with common root vertex v , but otherwise vertex disjoint.

It follows from Observation 6, that if we ignore the blue phases of the E -process, then the resulting red phases describe a continuous simple random walk $W_u(t_R)$ on the graph G . Each step t_R of the walk W_u corresponds to some step $s > t_R$ in the E -process. From Observation 6 it also follows that, if X starts at u , then W_u also starts at vertex u .

At step t of the E -process, we have $t = t_R + t_B$, where t_R , t_B are the (unknown) number of red and blue edge transitions. One thing is certain however; the length of the blue walk can be at most the number of edges m of G . This is formalized in the next observation.

Observation 8 *Let $W_u(t_R)$ be a simple random walk on the graph G defined by the red phase of the E -process, and let $X_u(t)$ be the walk defined by the E -process. Then $t_R < t < t_R + m$.*

2.2 Cover time of the E -process

Lemma 9 *Let \mathcal{W}_u be a random walk starting from u in G . Let S be a set of vertices of G of size s . Let*

$$d(S) = o(m/\log n),$$

where $d(S)$ be the sum of the degrees of the vertices in S . Let

$$t = \Omega(m/s(1 - \lambda_{\max})),$$

then

$$\Pr(S \text{ is unvisited by } \mathcal{W}_u \text{ at step } t) = O\left(e^{-td(S)(1-\lambda_{\max})/14m}\right).$$

Proof Contract S to a single vertex $\gamma = \gamma(S)$, retaining all resulting loops and parallel edges. Denote the resulting graph by Γ . Let $|S| = s$.

For $\lambda \leq 1$, $\lambda \leq e^{-(1-\lambda)}$. It follows from (2), for given u, x that

$$|P_u^t(x) - \pi_x| \leq \Delta^{1/2} e^{-(1-\lambda_{\max})t}, \tag{9}$$

where Δ is the maximum degree in G or Γ as appropriate. In either case, $\Delta \leq 2m = O(n^2)$. Let

$$T = K \log n / (1 - \lambda_{\max}),$$

where $K \geq 6$. As there are at most n^2 pairs u, x , then using (9)

$$\sum_{u,x} |P_u^t(x) - \pi_x| \leq n^2 \Delta^{1/2} e^{-T(1-\lambda_{\max})} = O(1/n^3).$$

Thus T is a mixing time satisfying (6) in both G and Γ . Also, from Corollary 5 we have

$$\mathbf{E}_\pi(H_S) \leq \frac{2m}{d(S)(1 - \lambda_{\max})}.$$

For $u \notin S$ let \mathcal{W}_u be a walk starting from u in G , and let $\widehat{\mathcal{W}}_u$ be the equivalent walk starting in Γ . Provided \mathcal{W}_u does not visit S in t steps, (the event $\mathbf{A}_t(S, G)$), then $\widehat{\mathcal{W}}_u$ does not visit γ (the event $A_t(\gamma, \Gamma)$), and the walks have the same probabilities. Thus

$$\mathbf{Pr}(\mathbf{A}_t(S, G)) = \mathbf{Pr}(\mathbf{A}_t(\gamma, \Gamma)).$$

From Lemma 4 we have

$$\mathbf{Pr}(\mathbf{A}_t(\gamma)) \leq \exp(-\lfloor t/(T + 3\mathbf{E}_{\hat{\pi}}(H_\gamma)) \rfloor).$$

Let T_Γ be a mixing time of the random walk on Γ satisfying (6). From (9), and the conditions on $t, d(S)$ given in the lemma, we have that $T_\Gamma = o(m/d(S)(1 - \lambda))$, and thus

$$T + 3\mathbf{E}_{\hat{\pi}}(H_\gamma) \leq \frac{7m}{d(S)(1 - \lambda_{\max})}.$$

We have the result that

$$\mathbf{Pr}(\mathbf{A}_t(S, G)) \leq \exp\left(-t \frac{d(S)(1 - \lambda_{\max})}{14m}\right).$$

□

Lemma 10 *Let G be a graph of maximum degree Δ . Let $\beta(s, v)$ be the number of connected edge induced subgraphs of size s rooted at vertex v in G . Then*

$$\beta(s, v) \leq 2^{s\Delta}.$$

Proof We make a crude estimate for $\beta(s, v)$ by building a digraph H_v in a breadth first manner as follows. Initially $H_v = \emptyset$ and all adjacent edges of v in G are labeled unvisited. Mark v as processed and add it to H_v . For each edge incident with v , we label it as retained or excluded. Starting from v there are $d(v)$ unvisited edges, and so at most $2^{d(v)}$ choices for the subset of edges incident with v to retain. We process each retained edge (v, u) in increasing endpoint label order. Mark u as processed and add the retained edge (v, u) to H_v . There are at most $2^{d(u)-1}$ choices for labels (retained, excluded) of any unvisited edges incident with u .

Thus we fan out from v in a breadth-first manner using only retained edges, (u, w) . We add w to H_v , and also any retained edges (x, w) , where x was processed earlier than w . In general there are some number of retained and excluded edges incident with w in G , resulting from processing earlier vertices; and the remaining at most $(d(w) - 1)$ edges are unvisited. We continue until H_v has s processed vertices, and the choices at these vertices have been evaluated. The s processed vertices of H_v and any retained edges between them defines a connected subgraph of size s rooted at v , and every subgraph of size s rooted at v is found by this construction. □

Lemma 11 *Let G be an ℓ -good graph of minimum degree δ and maximum degree Δ . With probability $1 - O(n^{-3})$, after*

$$\tau^* = O\left(m \left(1 + \frac{\Delta \log n}{\delta \min(\ell, \log n)(1 - \lambda_{\max})}\right)\right)$$

steps of the E -process, no vertex of G remains unvisited. The value of τ^* is independent of the choice of rule \mathcal{A} used by the process.

In particular, if G has constant maximum degree, there exists a constant $B > 0$ such that

$$\tau^* = Bn[1 + (\log n)/\min(\ell, \log n)(1 - \lambda_{\max})].$$

Proof Let S_v^* be the maximal connected even degree blue subgraph rooted at v , as described in Observation 7. Let S_v be any connected subgraph of S_v^* of size

$$s = \min(\ell, \log n),$$

rooted at v . By Lemma 10, there are at most $2^{\Delta s}$ such possible subgraphs.

For a random walk \mathcal{W}_u starting from vertex u , let $P(s, t)$ be the probability that at step t there exists an unvisited connected subgraph of size s rooted at some vertex v . Thus using Lemmas 9 and 10

$$P(s, t) \leq n2^{\Delta s} e^{-t \frac{d(S)(1-\lambda_{\max})}{14m}}.$$

As $s = \min(\ell, \log n)$, on choosing

$$t^* = (\Delta + 7) \log n \frac{14m}{\delta s(1 - \lambda_{\max})},$$

where $\delta \geq 2$ is minimum degree, we find that

$$P(s, t^*) = O(1/n^3). \tag{10}$$

From Observation 8, the length of the E -process walk on unvisited edges is at most m , the number of edges of G , and the step $\tau^* = \tau(t^*)$ in the E -process corresponding to the step t^* in the red phase random walk \mathcal{W}_u is bounded by $\tau^* \leq m + t^*$. In particular, if Δ is constant then $m = cn$, and

$$\tau^* \leq m + t^* = B(n + (n \log n)/(\min(\ell, \log n)(1 - \lambda_{\max}))).$$

Suppose some vertex v is unvisited at τ^* . Then a blue (unvisited) edge induced subgraph S_v^* rooted at v exists at τ^* . However, from (10), **whp** any $S_v \subseteq S_v^*$ of size s , contains a vertex z already visited by $\mathcal{W}(t^*)$. Suppose this visit occurs at $t \leq t^*$, but that, at step t^* , some edges incident with z are unvisited, a necessary condition for $z \in S_v^*$. On arriving at z , the E -process completes the exploration of all edges incident with z , after which the random walk $\mathcal{W}(t)$ continues up to step t^* . Thus at τ^* all edges adjacent to z are red, which is a contradiction. \square

3 Discussion and examples

3.1 Lower bound cover time for weighted random walks

For an introduction to properties of weighted random walks see [1]. The following proof that the cover time of any weighted random walk is $\Omega(n \log n)$, is due to T. Radzik [15].

For any vertex u , the expected first return time $\mathbf{E}T_u^+$ to u is $\mathbf{E}T_u^+ = 1/\pi(u)$.

The commute time $K(u, v)$ between vertices u and v , is the expected time taken to go from vertex u to vertex v and then back to vertex u . Formally, $K(u, v) = \mathbf{E}_u T_v + \mathbf{E}_v T_u$. Any walk starting

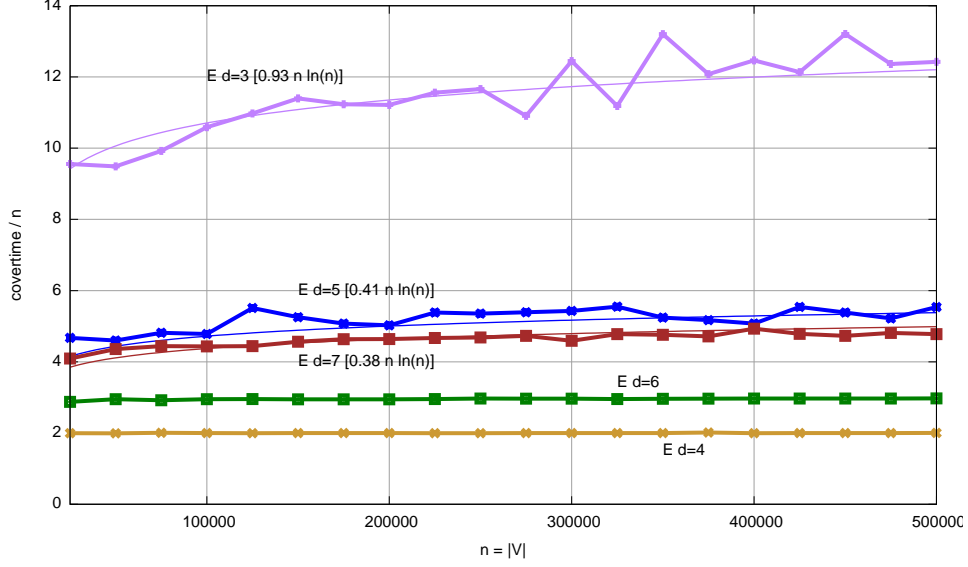


Figure 1: Normalised cover time of E -process as function of size and degree d

from u either visits v on the way back to u or it does not. Thus $\mathbf{E}T_u^+$ is at most the commute time $K(u, v)$ between u and v .

Let S be the subset of vertices with $\pi(u) \leq 2/n$. Thus $|S| \geq n/2$. This follows because $\sum_{u \in V} \pi(u) = 1$. As $\mathbf{E}T_u^+ = 1/\pi(u)$, it follows that for $u \in S$, $\mathbf{E}T_u^+ \geq n/2$.

Let $K_S = \min_{i,j \in S} K(i, j)$ then, $K_S \geq \mathbf{E}T_u^+ \geq n/2$. From [11], we have the lower bound that

$$C_G \geq (\max_{S \subseteq V} K_S \log |S|)/2 \geq (n/4) \log(n/2).$$

3.2 Proof of Corollary 2

Random r -regular graphs, \mathcal{G}_r , with $r \geq 4$ even, are an example of a class of graphs for which (**whp**) $C_G(E\text{-process}) = O(n)$. To establish this let \mathcal{G}'_r be the subset of \mathcal{G}_r with the following properties.

- (P1) G is connected, and the second eigenvalue of the adjacency matrix of G is at most $2\sqrt{r-1} + \varepsilon$, where $\varepsilon > 0$ is an arbitrarily small positive constant.
- (P2) Let $s = O(\log n)$, and let $a = \lceil 2s(\log re)/\log n \rceil$. No set of vertices S of size s induces more than $s + a$ edges. In particular, for $s \leq (\log n)/(2 \log re)$ no set of vertices S of size s induces more than s edges.

Lemma 12 *Let $\mathcal{G}'_r \subseteq \mathcal{G}_r$ be the r -regular graphs satisfying (P1), (P2). Then $|\mathcal{G}'_r| \sim |\mathcal{G}_r|$.*

Proof Friedman [10], shows the deep result that (P1) holds **whp** for random regular graphs. That (P2) holds **whp** is straightforward to establish. \square

Proof of Corollary 2. Let $\ell = \epsilon \log n$ for some $\epsilon > 0$. Property (P2) implies the graph is ℓ -good as follows. For any vertex v of the graph G , let U^* be the smallest non-trivial connected, even degree, vertex induced subgraph rooted at v . As $r \geq 4$, this subgraph contains at least two cycles.

Let $|U^*| = k$, then U^* induces at least $k + 1$ edges. By property (P2), no subgraph on $s = \epsilon \log n$ vertices with $\epsilon = 1/(2 \log re)$ induces more than s edges, and we conclude that $|U^*| > s$.

3.3 Removing the even degree constraint?

The only place in the proofs where the even degree condition matters is the proof of Observation 6, that the walk on unvisited edges terminates at its start vertex. How important is the even degree constraint?

We consider the experimental evidence for the performance of the E -process on both even degree, and odd degree graphs. In our experiments unvisited edges are chosen uniformly at random. We generated graphs of size up to half a million vertices, using the random regular graph generator from the *NetworkX* package (<http://networkx.lanl.gov/>) for the programming language Python. This package implements the Steger/Wormald approach, see [16]. We used Python's built-in random number generator which is based upon the Mersenne Twister. Each data point is the average of five actual experiments.

In Figure 1 we plot the normalised cover time of the E -process, in the case where the choice of unvisited edges is random. The *normalised* cover time is the actual cover time divided by n , as a function of n . Thus, linear functions of n appear flat etc. The labeling on the graphs is as follows: The first letter indicates an E -process, and this is followed by the degree $d = r$ of the graph. In the case where the plot appears to be non-linear, a curve of the form $c \log n$, is drawn behind the normalised experimental data, and labeled $[cn \ln(n)]$. The constant c used to draw the curve was determined by inspection.

It would appear the plots for even degrees 4 and 6 are constant, i.e. the cover time is $O(n)$. On the basis of experimental evidence, the normalised cover time of 3-regular graphs is $\omega(n)$; see Figure 1. This $\omega(n)$ growth appears to be $0.93n \log n$. For degrees 5 and 7 the plot also appears to grow logarithmically. We note, however, that it is notoriously difficult to quantify such growth on the basis of finite n , and we make no claims other than to present our experiments.

References

- [1] D. Aldous, J. Fill. *Reversible Markov Chains and Random Walks on Graphs*, 2001. <http://stat-www.berkeley.edu/users/aldous/RWG/book.html>
- [2] C. Avin and B. Krishnamachari. *The power of choice in random walks: An empirical study*. Proceedings of 9th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, (MSWiM-06), 219-228, (2006).
- [3] E. Bampas, L. Gąsieniec, N. Hanusse, D. Ilcinkas, R. Klasing, A. Kosowski. *Euler Tour Lock-in Problem in the Rotor-Router Model. I choose pointers and you choose port numbers*. Proceedings of DISC 2009, 423-435.
- [4] V. Bourassa and F. Holt, *SWAN: Small-world wide area networks*. Proceedings of International Conference on Advances in Infrastructure (SSGRR 2003w), L'Aquila, Italy, 2003, paper 64.
- [5] C. Cooper, M. Dyer, C. Greenhill. *Sampling regular graphs and a peer to peer network*. Combinatorics Probability and Computing 16(4) 557-593 (2007).

- [6] C. Cooper, M. Dyer, A. Handley. *The Flip Markov Chain and a Randomising P2P Protocol*. Proc. of PODC 2009. 141-150, (2009).
- [7] J. Cooper, B. Doerr, T. Friedrich, and J. Spencer. *Deterministic random walks on regular trees*. Proceedings of 19th ACM-SIAM Symposium on Discrete Algorithms (SODA'08), 766–772, (2008).
- [8] T. Feder, A. Guetz, M. Mihail, and A. Saberi. *A local switch Markov chain on given degree graphs with application in connectivity of peer-to-peer networks*. Proc. of FOCS, 2006.
- [9] U. Feige, *A tight lower bound for the cover time of random walks on graphs*. Random Structures and Algorithms 6 433–438 (1995).
- [10] J. Friedman. *A proof of Alon's second eigenvalue conjecture*. Proceedings of 35th ACM Symposium on Theory of Computing, 720–724 (2003).
- [11] J. Kahn, J. H. Kim, L. Lovasz, and V. H. Vu. *The cover time, the blanket time, and the Matthews bound* Proc. of FOCS'00, 467-475, (2000).
- [12] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica, 8, 261–277, (1988).
- [13] P. Mählmann and C. Schindelhauer. *Peer-to-peer networks based on random transformations of connected regular undirected graphs*. Proc. of SPAA '05, 155-164 (2005).
- [14] T. Orenshtein and I. Shinkar. *Greedy random walk*. [arXiv:1101.5711v3](https://arxiv.org/abs/1101.5711v3) [math.PR]
- [15] T. Radzik. Private communication. (2010).
- [16] A. Steger and N.C. Wormald. *Generating random regular graphs quickly*. Combinatorics, Probability & Computing 8(4): 377-396 (1999).
- [17] V. Yanovski, I. A. Wagner, and A. M. Bruckstein. *A distributed ant algorithm for efficiently patrolling a network*. Algorithmica, 37:165–186, 2003.